

Timeliness of Emergency Department Diagnoses for Syndromic Surveillance

Debbie Travers, PhD, RN; Clifton Barnett, MSIS; Amy Ising, MSIS; Anna Waller, ScD
Emergency Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC

Emergency Department (ED) data are key components of syndromic surveillance systems. While diagnosis data are widely available in electronic form from EDs and often used as a source of clinical data for syndromic surveillance, our previous survey of North Carolina EDs found that the data were not available in a timely manner for early detection. The purpose of this study was to measure the time of availability of participating EDs' diagnosis data in a state-based syndromic surveillance system. We found that a majority of the ED visits transmitted to the state surveillance system for 12/1/05 did not have a diagnosis until more than a week after the visit. Reasons for the lack of timely transmission of diagnoses included coding problems, logistical issues and the lack of IT personnel at smaller hospitals.

INTRODUCTION

Early detection of disease outbreaks can decrease patient morbidity and mortality and minimize the spread of diseases due to emerging infectious diseases (e.g., avian flu), bioterrorism (e.g., anthrax) and vaccine-preventable diseases (e.g., varicella). One approach to early detection is syndromic surveillance, in which electronic symptom data are captured early in the course of illness, and analyzed for signals that might indicate an outbreak requiring public health investigation and response.¹ Increasingly, public health surveillance systems are utilizing secondary data from electronic health records because the data can be timely, comprehensive and population-based. Emergency department (ED) records are a frequent data source for these systems.²⁻³ ED data have been shown to detect outbreaks 1-2 weeks earlier than through traditional public health reporting channels.^{2,4-6} Though there is no formal standard for how soon after an ED visit or other health system encounter data should be available for early detection, most surveillance systems aim for near real-time data within hours, not days or weeks.

Several challenges to using ED data for surveillance have been identified. Many EDs still document patient symptoms manually and even when the data are electronic, they may be entered in free text form instead of using standardized terms. Timeliness is also a concern; while some ED data elements are entered into electronic systems at the start of the ED visit, other elements are added hours, days or even weeks later.

The most widely used ED data element for syndromic surveillance is the chief complaint (CC), which is a brief description of the patient's primary symptom(s). The CC is available in electronic form at most EDs, and is typically entered in the first few minutes of the ED visit. However, there are no national standards for CC, which is often recorded in free text and includes misspellings, abbreviations, acronyms and other lexical and semantic variants that are difficult to aggregate into symptom clusters.^{2,6-9}

Another ED data element used for surveillance is the final diagnosis, which is widely available in electronic form and is standardized using the International Classifications of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM).^{2,7,10} All EDs generate electronic ICD-9-CM diagnoses since this is required for billing purposes.^{9,11} There is, however, some evidence that diagnosis data are not always available in a timely manner. In contrast to CC data which are generally entered into ED information systems by clinicians in real time, ICD-9-CM diagnoses are often entered into the system by coders well after the ED visit.¹² In a 2003 study of regional surveillance systems in North Carolina and Washington, we found that over half of the EDs did not have electronic diagnosis data until 1 week or more after the ED visit.⁹ Sources of ICD-9-CM data may vary, which may influence the quality of the data. Traditionally, diagnoses have been assigned to ED visits by trained coders who are employed by the hospital and/or physician professional group. The primary purpose of the coding is billing, as opposed to secondary uses such as surveillance. Recently, emergency department information systems (EDIS) have come on the market that allow for diagnosis entry by clinicians. The systems typically include drop-down boxes with text that correspond to ICD-9-CM codes; clinicians can then select a "clinical impression" at the end of the ED visit.

The North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) system is a state-based surveillance system that utilizes ED, poison center, prehospital and veterinary laboratory data.¹³ The system began as a pilot project in 1999, with three EDs sending data. The state of NC recently mandated that ED data be submitted to the state for public health surveillance via NC DETECT, and all EDs now participate or are preparing to participate. NC DETECT currently receives data from 75 of the 113 (66%) EDs in the state, and is expected to include data from all NC

EDs by the end of 2006. To date, the database includes over 2.5 million visits, with some hospitals reporting historical data from 2000 to present. The majority of the participating EDs send data to a third party data aggregator via secure hypertext transfer protocol, who then sends it on to NC DETECT twice daily. The data aggregator excludes visits for patients not treated in the ED (e.g., sent directly to labor and delivery) and visits missing required basic visit data. Some hospitals send data directly to NC DETECT via encrypted file transfer protocol, but this process is being phased out. Upon receipt of data by NC DETECT, duplicate visits are deleted and the data are then transformed and loaded into a MS SQL Server (Redmond, WA) database and are used for a variety of ad hoc and standardized reports.

The syndromic surveillance component of NC DETECT monitors the ED data daily for 7 syndromes: botulism, fever/rash, gastrointestinal, influenza-like illness, meningococcal, neurological and respiratory. The syndrome reports are generated through locally developed syndrome queries written in standard query language. Currently, the following data elements are used to identify positive cases for the syndrome reports: CC and, when available, temperature and triage note (a free text history of present illness that expands on the CC). While triage note has been shown to improve the identification of positive cases, currently only 8 NC DETECT hospitals send these notes.¹⁴ Based on the results of our 2003 survey of EDs, we elected not to include diagnosis data in the syndrome queries, since a majority of NC EDs reported that they could not provide ICD-9-CM codes to NC DETECT in less than one week from the visit date.

Table 1- NC DETECT- Selected Data Elements

Data Elements	Required With Initial Transmission
Unique ID Visit date, arrival time Facility ID Chief complaint	Yes
Triage note Diagnosis code(s) (1-11) External cause of injury codes (1-5) Procedure codes (1-10) Vital signs Zip code	No

Select NC DETECT data elements are required in order to accept a visit into the system. Other data can be added in subsequent record updates and may be less timely for early detection purposes. The NC DETECT system time stamps each record update.

Table 1 shows the NC DETECT data elements that may be used for surveillance, and indicates which ones are required in the initial transmission. Of the NC DETECT data elements that include clinical information potentially useful for syndromic surveillance, only the CC is available with the initial transmission of data. Triage notes are optional and if sent, may be transmitted later. A minimum of one (primary), and up to 10, diagnoses are required by NC DETECT, but these may be added after the initial transmission. The purpose of this study was to measure the time of availability of ED ICD-9-CM data to NC DETECT for syndromic surveillance.

METHODS

In this descriptive study, we prospectively measured the time of availability of electronic ICD-9-CM codes in the NC DETECT database for all ED visits on 12/1/05. Included were all visits transmitted to NC DETECT from the 57 EDs that were actively participating in the system as of 12/1/05. The visit was considered to have available diagnosis data if it included one or more valid ICD-9-CM codes.

We used time stamp data (a record of the time that all data are received, including initial visit data and any record updates sent later) to measure the number of 12/1/05 visits that had 1+ diagnosis code at intervals, starting with noon on 12/2 (12 hours), and including 24 and 36 hours, 2-7 days, and 1, 2, 3, 4, 6, 8, 10 and 12 weeks. Other data included time intervals for receipt of CC and basic visit data. We also sent email surveys to all NC DETECT hospitals asking about their source of ICD-9-CM data, and their procedures for extraction and transmission of diagnosis data. A final source of data was NC DETECT data quality logs, in which problems with data are tracked from identification to resolution.

RESULTS

Table 2- Number of ICD-9-CM Codes per Visit for 12/1/05

N= 5,319 ED visits, N= 16,118 diagnoses

Diagnosis Codes Per Visit	Visit Count
0	727
1	1180
2	1135
3	701
4	435
5	273
6	202
7	129
8	96
9	81
10	112
11 or more	248

There were 5,319 ED visits from 57 hospitals for the study enrollment date, 12/1/05. The number of visits per ED ranged from 7 to 284, with a mean of 93 and a median of 71. Table 2 shows the number of diagnoses per visit, as of 12 weeks after 12/1/05. As of 12 weeks after the study enrollment date, 4,592 (86%) of the 5,319 visits had at least one diagnosis, while 727 (14%) had yet to have any diagnoses sent. A total of 16,118 diagnosis codes had been transmitted to NC DETECT for the study date. For the 4,592 visits with at least one diagnosis code, there were 3.5 diagnoses per visit.

**Figure 1-Timeliness of Diagnosis, CC & Visit Data
N= 5,319 ED visits**

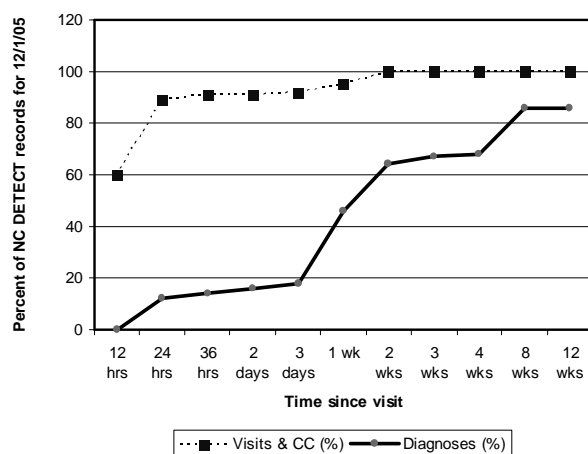


Figure 1 shows the timeliness of basic visit data transmitted to NC DETECT at major time intervals, with the percent of the 5,319 visits that included CC and diagnosis data. All visits contained a CC which was required for the initial transmission of the record to NC DETECT. However, at the 12-hour mark (noon on 12/2/05), only 60% of the visits and CCs for 12/1/05 had been transmitted to NC DETECT. At 24 hours, 89% of the visits/CCs had been sent and at 2 weeks 100% were available in NC DETECT.

There were almost no diagnoses for 12/1/05 visits available in NC DETECT at noon the following day. At the 24 hour interval, only 12% of the visits had at least one diagnosis code. The majority of the visits did not have a diagnosis until more than a week after the visit. By the 2 week interval, 64% of the visits had diagnosis data available for surveillance, and at 12 weeks 86% of the visits had one or more diagnoses.

The availability of at least one diagnosis code for each visit from the 57 NC DETECT hospitals ranged from 12 hours to greater than 12 weeks. At 12 hours after 12/1/05 (noon on 12/2), only two hospitals had sent any diagnoses: one hospital had diagnoses for 1% of the visits and another for 3%. By the 24 hour

mark, 5 hospitals had 100% of their visits coded with at least one diagnosis. At the other end of the spectrum, 7 hospitals had diagnoses for less than two-thirds of their visits at the 12 week mark. Table 3 shows the EDs with the most and least available data by time, and includes the number of ED visits sent to NC DETECT by each site for 12/1/05.

**Table 3- Selected NC DETECT EDs:
Most and Least Timely Data**

ED	% visits with 1+ ICD-9-CM code(s)	Time Interval	# of ED Visits for 12/1/05
A	100%	24 hrs	217
B	100%	24 hrs	140
C	100%	24 hrs	56
D	100%	24 hrs	61
E	100%	24 hrs	110
F	62%	12 wks	34
G	50%	12 wks	44
H	35%	12 wks	23
I	28%	12 wks	169
J	20%	12 wks	229
K	11%	12 wks	96
L	9%	12 wks	54

33 (58%) of the 57 participating hospitals responded to the email survey with information about the diagnosis data they send to NC DETECT. They reported two different sources of ED diagnosis data: coders and clinicians. All 33 EDs send data entered by coders into hospital/billing systems. Six EDs also send a single, primary diagnosis entered by a clinician from the EDIS. Three sites explained that they do have emergency department information systems but do not pull diagnosis codes from them because they are either not stored in the EDIS at all, or the system is set up in such a way that it is not possible to extract them. Some users voiced concerns about the validity of clinician-entered diagnosis data. Others reported that some EDIS allow hospitals to add items to the drop-down diagnosis boxes that may not be valid ICD-9-CM codes.

The sites also reported on the timing of their diagnosis code data transmission to NC DETECT. 25 of the 33 sites send their ICD-9-CM codes as soon as the codes have been entered, whereas 8 transmit the codes in batch format (e.g., weekly, monthly).

The data quality logs revealed several problems that have resulted in problems in transmitting ICD-9-CM codes to NC DETECT. At one site, the procedure for extraction of ICD-9-CM codes from the billing system was changed due to a new hospital-wide IT initiative, and subsequent diagnosis data were only sent to NC DETECT for the 28% of ED patients who were admitted to an inpatient unit from the ED. At another hospital, there is only one coder

that codes ED visits and when that person is away from work, the coding is delayed. Other sites have limited IT staff and if there is a problem with data transmission on weekends/holidays, there is a delay to the next business day.

DISCUSSION

Diagnosis data are universally available from NC emergency departments. Studies have shown that ICD-9-CM data alone or in combination with CC data results in higher sensitivities and better agreement with expert reviewers than CC data alone for syndromic surveillance.^{7,12,15} However, in spite of the high value of diagnoses for syndromic surveillance, this study corroborated our earlier hospital survey results that indicated the majority of NC DETECT hospitals cannot send diagnosis data soon enough for timely, population-based syndromic surveillance. In this study, we conducted actual measurements of available data in our state-based syndromic surveillance system, an improvement over our previous study methods which relied on self-report by hospitals. We confirmed that fewer than half of the EDs sent diagnoses within one week of the visit, and that it took three weeks to get at least one diagnosis for two-thirds of the visits.

The major factors associated with delays in transmission of diagnosis data from EDs to NC DETECT include coding delays, other logistical issues and limited IT resources. Many of these problems were distributed across both small and large EDs. It is not uncommon for the coding of ED diagnoses to be completed several days or more after the ED visit. The hospital billing and EDIS systems from which NC DETECT data are extracted are diverse and many are not easy to extract data from. The NC DETECT team has developed customized processes to facilitate data extraction from some hospitals. We have also found the need to continue to monitor sites that have been deemed fully operational, since these extraction processes can be affected when hospitals change various IT processes and/or information systems over time. We have also found a lack of IT personnel to help with the NC DETECT system, at some sites, particularly at small hospitals. At these sites, data transmission delays are more common, e.g., data are batched or may not be transmitted on weekends.

These limitations on the timely transmission of data occur in spite of the NC state law mandating that EDs provide their available electronic data (including ICD-9-CM codes) to NC DETECT. There is no mechanism in place for enforcement of this mandate, and the project has limited resources to devote to data quality monitoring such we have conducted for this study, and for follow up with hospitals.

In this study we identified another delay that affects the timeliness of diagnosis data availability: the limitation of the timing of basic visit data transmission. While the goal of the NC DETECT system is to be a near-real-time system, in fact the data are transmitted from the third party data aggregator to NC DETECT twice daily. We found that only 60% of the ED visits for 12/1/05 had been transmitted to NC DETECT by noon on 12/2/05.

As more hospitals adopt EDIS, these systems may emerge as a more timely source of ED diagnoses. However some issues surrounding this source of data will need further investigation, including validity of diagnosis coding by clinicians, and the need for methods to extract the diagnosis data from the EDIS.

In the event that ICD-9-CM data were more widely available from EDs in a timely manner, there is still a question of whether or not ICD-9-CM are valid for public health surveillance. Potential limitations of diagnosis data for surveillance include coding variability, potential coding errors, and the challenge of determining a definitive diagnosis after a brief ED visit.¹⁶⁻¹⁷ Further study is needed to validate sensitivity, specificity and positive and negative predictive values of ICD-9-CM codes from ED visits used for syndromic surveillance. Diagnosis data in surveillance systems should also be validated to ensure that codes actually match the information in the medical record.¹⁸

Though the goal of syndromic surveillance is early detection of disease outbreaks, there is no standard for how soon after the ED visit that data should be available for analysis. While some systems aim for real-time data collection and others for near real-time, the definitions for these terms have yet to be established what real-time means, versus near real-time. More research is needed to quantify how timely existing surveillance systems actually are, and to correlate time data with outcomes.

Overall, other sources of NC DETECT data are more timely for near real-time syndromic surveillance than diagnosis data. In both this and our previous study, we found the CC to be the most timely and widely available source of data for surveillance from NC EDs.⁹ We have found that when available, triage note data are also timely and improve our ability to detect illnesses of interest five-fold, but the data are only available from a minority of NC DETECT hospitals.¹⁴

However, while overall, the diagnosis data in NC DETECT are not that timely, there are some notable exceptions. As shown in Table 3, five hospitals transmit diagnosis data within 12 hours of the visit. These hospitals have EDIS in which clinicians enter a primary diagnosis at the end of the ED visit. Some of these hospitals do not provide triage notes to NC

DETECT, so it may be useful to add diagnosis codes to our syndrome queries for these hospitals.

This study has provided us with the ability to identify problems and provide feedback regarding individual EDs' transmission of diagnosis data. We are also using the results to inform the NC DETECT data quality efforts, e.g., exploring automated monitoring of time availability of diagnosis data.

The limitations of this study include the fact that we took time measurements from only one date, 12/1/05. While it is possible that we have a biased sample, our technical staff have subsequently performed similar checks and confirmed these trends.

CONCLUSIONS

While timely ED diagnosis data are desirable for syndromic surveillance, we found that even in a state with a mandate requiring EDs to report data to the statewide public health surveillance system, that the data may not be available for early detection. The latency of diagnosis data is influenced by logistical issues and the fact that contributing EDs do not have dedicated personnel for this surveillance. This study illustrates the importance of evaluating the quality of the data used for surveillance.

ACKNOWLEDGEMENTS

Funding for NC DETECT is provided by the NC Division of Public Health. Thanks to Lorraine Waguespack for her contributions to this paper.

REFERENCES

1. Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V. Framework for evaluating public health surveillance systems for early detection of outbreaks. *MMWR*. 2004;53(RR05):1-11.
2. Lober WB, Karras BT, Wagner MM et al. Roundtable on bioterrorism detection: information system-based surveillance. *J Am Med Inform Assoc*. 2002;9(2):105-15.
3. Teich JM, Wagner MM, Mackenzie CF, Schafer KO. The informatics response in disaster, terrorism and war. *J Am Med Inform Assoc*. 2002;9(2):97-104.
4. Tsui FC, Wagner MM, Dato V, Chang CC. The value of ICD-9 coded chief complaints for detection of epidemics. *Proc AMIA Symp*. 2001;711-715.
5. Heffernan R, Mostashari F, Das D et al. New York City syndromic surveillance systems. *MMWR*. 2004;24 (53S):23-7.
6. Wagner MM, Espino J, Tsui FC et al. Syndrome and outbreak detection using chief-complaint data- experience of the Real-Time Outbreak and Disease Surveillance project. *MMWR*. 2004;24(53S): 28-31.
7. Fleischauer AT, Silk BJ, Schumacher M et al. The validity of chief complaint and discharge diagnosis in emergency-based syndromic surveillance. *Acad Emerg Med*. 2004;11(12):1262-7.
8. Shapiro AR. Taming variability in free text: Application to health surveillance. *MMWR*. 2004; 24(53S):95-100.
9. Travers DA, Waller A, Haas SW, Lober WB, Beard C. Emergency Department data for bioterrorism surveillance: Electronic data availability, timeliness, sources and standards. *AMIA Annu Symp Proc*. 2003;:664-8
10. U.S. Department of Health and Human Services (USDHHS). International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). 5th ed. Washington: Author; 1999.
11. National Center for Injury Prevention and Control (US). Data elements for emergency department systems, release 1.0. Atlanta, GA: Centers for Disease Control; 1997.
12. Beitel AJ, Olson KL, Reis BY, Mandl KD. Use of emergency department chief complaint and diagnostic codes for identifying respiratory illness in a pediatric population. *Peds Emerg Care*. 2004;20(6):355-360.
13. Ising A, Waller A, Ahmad F et al. North Carolina Emergency Department Database web-based portal: timely, convenient access to ED data for syndromic surveillance. *Proc Nat Syn Surv Conf*. 2004.
14. Ising A, Travers D, MacFarquhar J, Kipp A, Waller A. Triage note in emergency department-based syndromic surveillance. *Adv Dis Surv* (in press).
15. Reis BY, Mandl KD. Syndromic surveillance: The effects of syndrome grouping on model accuracy and outbreak detection. *Ann Emerg Med*. 2004;44:235-241.
16. Chang HH, Glas RI, Smith PF et al., Disease burden and risk factors for hospitalizations associated with rotavirus infection among children in New York State, 1989 through 2000. *Pediatr Infect Dis J*. 2003;22:808-14.
17. Rosamond WD, Chambliss LE, Sorlie PD et al. Trends in the sensitivity, positive predictive value, false-positive rate, and comparability ratio of hospital discharge diagnosis codes for acute myocardial infarction in four US communities, 1987-2000. *Am J Epidemiol*. 2004;160:1137-46.
18. German RR. Sensitivity and predictive value positive measurements for public health surveillance systems. *Epidemiology*. 2000;11(6):720-27.